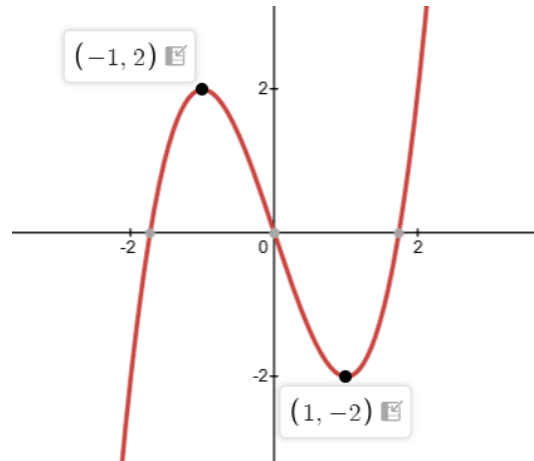


數學示例：變分法

微積分學的一個重要應用是求連續實值函數的「局部最大／最小化點」(local maximizing / minimising point)，即在局部範圍內令函數達到最大／最小值的點。舉例說，考慮函數 $y(x) = x^3 - 3x$; $x \in (-\infty, \infty)$ ¹，下圖展示這個函數的圖象：



學過微積分的讀者都知道，上述函數的「局部最大／最小化點」必須是以下方程的根：

$$Dy(x) = 0 \quad (1)$$

以上圖所示函數為例，這個函數的 1 階導數是 $Dy(x) = 3x^2 - 3$ ，而方程 $3x^2 - 3 = 0$ 的根是 $x = \pm 1$ ，由此可知 $x = \pm 1$ 可能是這個函數的「局部最大／最小化點」。

惟請注意，(1) 只是某點作為「局部最大／最小化點」的「必要條件」而非「充分條件」，因為滿足 (1) 的點可能是「拐點」(inflecting point) 而非「局部最大／最小化點」，因此在求得 (1) 的根後，我們還要判斷這些根所對應的點是「局部最大化點」、「局部最小化點」還是「拐點」。為此，可以應用其

¹為符合「變分法」一般採用的符號，並且考慮到數學上通常使用 x 和 y 坐標表示函數的圖象，本文用 y (而非 f) 代表函數。

他數學工具，例如描繪上述函數的圖象或者求它的 2 階導數。不過，在很多情況下，如果有關函數是來自一個應用問題，單憑問題的性質，便可判斷有關駐點屬於哪一種類型。利用上述方法，可以求得 $x = -1$ 和 $x = 1$ 分別是上述函數的「局部最大化點」和「局部最小化點」，對應這兩點的「局部最大值」(local maximum) 和「局部最小值」(local minimum) 分別為 2 和 -2 。

本文介紹的變分法(calculus of variations) 可被看成上述微積分應用的推廣。但跟上述概念不同，變分法主要不是用來尋找能使某給定「函數」在某範圍內達到「局部最大／小值」的點(即「局部最大／小化點」)，而是尋找能使某給定「泛函」(functional) 在某範圍內達到「局部最大／小值」的函數(可稱為「局部最大／小化函數」)。前述「微積分應用」與「變分法」的區別可以概括為：前者的目的是求「(函數的) 點」，後者的目的則是求「函數」；前者的給定資料是「函數」，後者的給定資料則是「泛函」。

在「泛函分析」中，「泛函」有專門的定義，但在本文中，可以把「泛函」簡單理解為「函數的實值函數」，即把函數映射為實數的函數，而在「變分法」的理論中，這個泛函一般表現為以下定積分：

$$\int_{x=x_1}^{x=x_2} F(x, y(x), Dy(x)) \quad (2)$$

其中 F 代表某個以區間 $[x_1, x_2]$ 為定義域並且其值依賴於 x 、 $y(x)$ 和 $Dy(x)$ 的連續函數關係。請注意上式的值是實數，而這個值會隨著函數 $y(x)$ 而變化，所以確是一個把函數映射為實數的函數(即「泛函」)。在變分法中，有一個類似 (1) 的方程，此即以下的歐拉-拉格朗日微分方程(Euler-Lagrange differential equation)：

$$D([D_w F(u, v, w)]_{(u,v,w)=(x,y(x),Dy(x))}) - [D_v F(u, v, w)]_{(u,v,w)=(x,y(x),Dy(x))} = 0, \\ x \in [x_1, x_2], \quad y(x_1) = y_1, y(x_2) = y_2 \quad (3)$$

其中的 F 就是 (2) 中的函數 F 。請注意在上式中， u 、 v 和 w 是三元函數 $F(u, v, w)$ 的自變項，所以作用於這個函數的是偏導數算子 D_w 和 D_v ；但 $[D_w F(u, v, w)]_{(u,v,w)=(x,y(x),Dy(x))}$ 卻是以 x 作為自變項的一元函數，所以作用於這個函數的是普通的導數算子 D 。另請注意如同 (1) 的情況，(3) 只是某函數作為「局部最大／最小化函數」的「必要條件」而非「充分條件」，在求得滿足上述方程的函數後，我們還要應用其他數學工具，或者根據問題的性質，來判斷這個函數是「局部最大化函數」、「局部最小化函數」還是兩者皆不是。在以下的例子中，我們將集中討論如何求解「歐拉-拉格朗日(微分／差分) 方程」，並假設可用相關的工具判斷所求得解就是當前問題的答案。

以下提供一個簡單例子以說明上述概念，考慮平面上的兩點 (x_1, y_1) 和

(x_2, y_2) 以及連接這兩點的任意連續曲線 (包括直線)。現設我們要在這些曲線中尋找其長度為最短的曲線, 即「微分幾何」中所稱的「測地線」(geodesic)。根據直觀, 我們知道直線距離是平面上兩點之間的最短距離, 因此直線就是我們要求的測地線, 以下讓我們用變分法的術語表述和求解此一問題。

根據《數學示例：曲線的參數化》，平面上 (x_1, y_1) 和 (x_2, y_2) 之間的曲線的函數 $y(x)$ 可以寫成以下「參數化形式」(下式來自上述網頁的 (1))：

$$Z : [x_1, x_2] \rightarrow \mathbb{R}^2; \quad Z(x) = (x, y(x)) \quad (4)$$

應用上述網頁的 (7) 和 (6), 可以用下式求上述曲線在 (x_1, y_1) 和 (x_2, y_2) 之間一段的長度 (即「弧長」 arc length)²：

$$\int_{x=x_1}^{x=x_2} \sqrt{1 + (Dy(x))^2} \quad (5)$$

上式就是與「平面測地線問題」相關的泛函。把 (5) 與 (2) 比較, 可以看到我們有 $F(x, y(x), Dy(x)) = \sqrt{1 + (Dy(x))^2}$, 由此有函數 $F(u, v, w) = \sqrt{1 + w^2}$ 。把此一函數代入 (3) 中等號左端的 F , 可得到

$$\begin{aligned} & D \left(\left[D_w \left(\sqrt{1 + w^2} \right) \right]_{(u,v,w)=(x,y(x),Dy(x))} \right) - \left[D_v \left(\sqrt{1 + w^2} \right) \right]_{(u,v,w)=(x,y(x),Dy(x))} \\ &= D \left(\left[\frac{w}{\sqrt{1 + w^2}} \right]_{(u,v,w)=(x,y(x),Dy(x))} \right) - 0 \\ &= D \left(\frac{Dy(x)}{\sqrt{1 + (Dy(x))^2}} \right) \end{aligned}$$

由此可得「平面測地線問題」的「歐拉-拉格朗日微分方程」如下：

$$D \left(\frac{Dy(x)}{\sqrt{1 + (Dy(x))^2}} \right) = 0, \quad x \in [x_1, x_2], \quad y(x_1) = y_1, y(x_2) = y_2 \quad (6)$$

從上述方程, 容易得到

$$\frac{Dy(x)}{\sqrt{1 + (Dy(x))^2}} = c_1 \quad (7)$$

²根據上述網頁的 (7) 和 (6), 若某曲線的函數可寫成參數化形式 $Z(t) = (x(t), y(t))$, 則該曲線在 $t = t_0$ 和 $t = t_1$ 之間一段的弧長可用定積分 $\int_{t=t_0}^{t=t_1} \sqrt{(Dx(t))^2 + (Dy(t))^2}$ 計算, (5) 就是應用此定積分公式的結果。

其中 c_1 是任意常數。從上式可以解出 $Dy(x) = \pm\sqrt{\frac{c_1^2}{1-c_1^2}}$ ，把這個結果中的 $\pm\sqrt{\frac{c_1^2}{1-c_1^2}}$ 換成另一個任意常數 c_2 ，可得

$$Dy(x) = c_2 \quad (8)$$

對上式左右兩端進行不定積分運算，或把上式看成 1 階常係數非齊次微分方程，可求得上述方程的通解為

$$y(x) = c_2x + c_3 \quad (9)$$

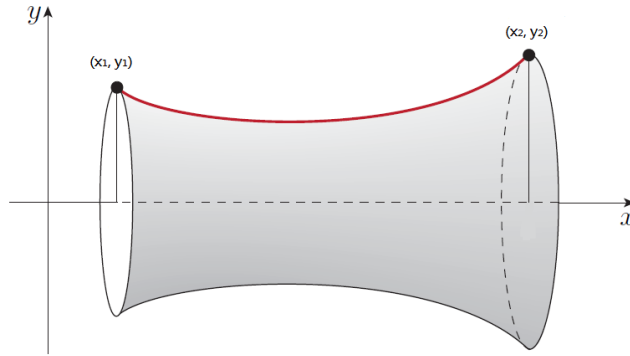
上式是直線的方程，至此我們看到直線確是平面上的測地線。把 (6) 中的邊界條件代入上式，便可解出上式中的任意常數 c_2 和 c_3 ，從而得到 (6) 的特解為

$$y(x) = \frac{y_2 - y_1}{x_2 - x_1}x + \frac{x_2y_1 - x_1y_2}{x_2 - x_1}, \quad x \in [x_1, x_2] \quad (10)$$

一般來說，方程 (3) 並不容易求解，但如果 (3) 中的 F 缺少了某個自變項 u 、 v 或 w ，可以把 (3) 簡化。特別地，如果 F 缺少了自變項 u ，那麼可以把 (3) 簡化成以下方程³：

$$\begin{aligned} Dy(x) \times [D_w F(v, w)]_{(v,w)=(y(x), Dy(x))} - F(y(x), Dy(x)) &= c_1, \\ y(x_1) &= y_1, y(x_2) = y_2 \end{aligned} \quad (11)$$

其中 c_1 是任意常數。舉例說，考慮平面上連接兩點 (x_1, y_1) 和 (x_2, y_2) 的任意連續曲線，並把該曲線繞 x 軸旋轉以形成一個「旋轉曲面」(surface of revolution)，如下圖所示：



現設我們要在這些曲線中尋找使得所形式的旋轉體面積為最小的曲線。根據微積分的知識，上圖所示旋轉曲面的面積可用以下公式計算：

$$\int_{x=x_1}^{x=x_2} 2\pi y(x) \sqrt{1 + (Dy(x))^2} \quad (12)$$

³讀者請自行驗證，如用下式處理「平面測地線問題」，可同樣推導出方程 (8)。

上式就是與「最小旋轉曲面問題」相關的泛函。把 (12) 與 (2) 比較 (略去 (2) 中的變項 x)，可以看到我們有 $F(y(x), Dy(x)) = 2\pi y(x)\sqrt{1 + (Dy(x))^2}$ ，由此有函數 $F(v, w) = 2\pi v\sqrt{1 + w^2}$ 。把此一函數代入 (11) 中等號左端的 F ，可得到

$$\begin{aligned}
& Dy(x) \times \left[D_w \left(2\pi v\sqrt{1 + w^2} \right) \right]_{(v,w)=(y(x), Dy(x))} - 2\pi y(x)\sqrt{1 + (Dy(x))^2} \\
&= Dy(x) \times \left[\frac{2\pi vw}{\sqrt{1 + w^2}} \right]_{(v,w)=(y(x), Dy(x))} - 2\pi y(x)\sqrt{1 + (Dy(x))^2} \\
&= \frac{2\pi y(x)(Dy(x))^2}{\sqrt{1 + (Dy(x))^2}} - 2\pi y(x)\sqrt{1 + (Dy(x))^2} \\
&= -\frac{2\pi y(x)}{\sqrt{1 + (Dy(x))^2}}
\end{aligned}$$

由此可得「最小旋轉曲面問題」的「歐拉-拉格朗日微分方程」如下 (在下式中，我們把 $-\frac{c_1}{2\pi}$ 換成另一任意常數 c_2)：

$$\frac{y(x)}{\sqrt{1 + (Dy(x))^2}} = c_2, \quad x \in [x_1, x_2], \quad y(x_1) = y_1, y(x_2) = y_2 \quad (13)$$

上述方程可以改寫成以下形式 (在下式中，我們把 $\pm c_2$ 換成另一任意常數 c_3)：

$$Dy(x) = \frac{1}{c_3} \times \sqrt{(y(x))^2 - c_3^2}, \quad x \in [x_1, x_2], \quad y(x_1) = y_1, y(x_2) = y_2 \quad (14)$$

請注意上式具有《數學示例：可分變項與可分核》中 (1) 的形式，其中 $\sqrt{(y(x))^2 - c_3^2}$ 可看成把函數 $h(x) = \sqrt{x^2 - c_3^2}$ 作用於 $y(x)$ 的結果。接著利用該網頁的 (5) 求解上述方程如下⁴：

$$\begin{aligned}
\left(\int \frac{1}{\sqrt{x^2 - c_3^2}} \right) \circ y(x) &= \int \frac{1}{c_3} \\
\operatorname{arccosh} \left(\frac{x}{c_3} \right) \circ y(x) &= \frac{x}{c_3} + c_4 \\
\operatorname{arccosh} \left(\frac{y(x)}{c_3} \right) &= \frac{x}{c_3} + c_4 \\
y(x) &= c_3 \cosh \left(\frac{x}{c_3} + c_4 \right) \quad (15)
\end{aligned}$$

⁴根據一般的「積分表」， $\int \frac{1}{\sqrt{x^2 - c_3^2}} = \ln \left(\frac{x}{c_3} + \sqrt{\frac{x^2 - c_3^2}{c_3^2}} \right)$ (略去任意常數)；但可以證明， $\ln \left(\frac{x}{c_3} + \sqrt{\frac{x^2 - c_3^2}{c_3^2}} \right) = \operatorname{arccosh} \left(\frac{x}{c_3} \right)$ 。

上述結果顯示，「最小旋轉曲面問題」的解涉及「反雙曲函數」。把 (13) 中的邊界條件代入上式，原則上可解出上式中的任意常數 c_3 和 c_4 ，從而得到 (13) 的特解。

在傳統的變分法中，(2) 中的函數 F 是連續函數，因此這種變分法可稱為**連續變分法**(continuous calculus of variations)。但我們也可以考慮**離散變分法**(discrete calculus of variations)，這種變分法的泛函一般表現為以下定和分：

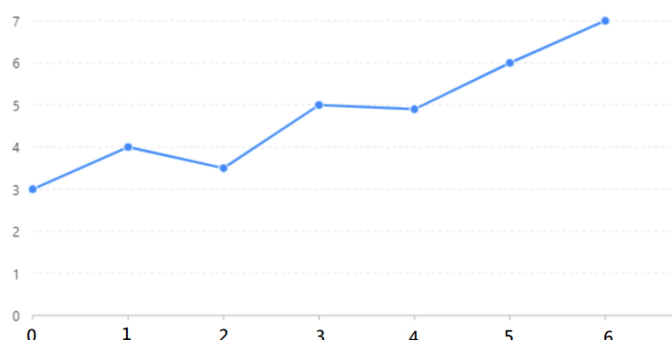
$$\sum_{x=x_1+1}^{x=x_2} F(x, y(x), \Delta y(x-1)) \quad (16)$$

上式跟 (2) 的區別在於，上式中的 F 是以離散整數集合 $\{x_1, x_1+1, \dots, x_2-1, x_2\}$ 為定義域並且其值依賴於 x 、 $y(x)$ 和 $\Delta y(x-1)$ 的離散函數關係，而且這個泛函的值是通過對 F 的離散值求定和分而得。不過，上式跟 (2) 一樣，也是一個把函數映射為實數的函數 (即「泛函」)。對應於 (16) 這個泛函，我們有以下這個**歐拉-拉格朗日差分方程**(Euler-Lagrange difference equation)：

$$\Delta([D_w F(u, v, w)]_{(u,v,w)=(x,y(x),\Delta y(x-1))}) - [D_v F(u, v, w)]_{(u,v,w)=(x,y(x),\Delta y(x-1))} = 0, \\ x \in \{(x_1), x_1+1, \dots, x_2-1, (x_2)\}, y(x_1) = y_1, y(x_2) = y_2 \quad (17)$$

其中的 F 就是 (16) 中的函數 F 。

舉例說，考慮平面上的兩點 (x_1, y_1) 和 (x_2, y_2) 以及連接這兩點的折線 (polygonal chain)，這裡的「折線」是指由一條或多條直線組成的線段。為簡單起見，以下假設組成折線的每條直線有相等的橫向距離 1，因此共有 $x_2 - x_1$ 條直線，例見下圖：



上圖展示一條連接 $(0, 3)$ 與 $(6, 7)$ 這兩點的藍色折線。組成此折線的直線的端點 (即上圖中的藍色點) 的 x 坐標是 0、...、6，這些直線的橫向距離全都是 1，因此共有 $6 - 0 = 6$ 條直線。如果我們把這些端點的 y 坐標與

x 坐標的關係處理成離散函數 $y(x)$ ，那麼每個其 x 坐標為 n 的端點的 y 坐標可以寫成 $y(n)$ ，而上圖中第 n 條直線 ($1 \leq n \leq 6$) 的縱向距離就是 $y(n) - y(n-1)$ ，亦即 $\Delta y(n-1)$ 。請注意由於每條折線的橫向距離都是 1，因此 $\Delta y(n-1)$ 其實也等於第 n 條折線的斜率。

現設我們要在上述折線中尋找其長度最短的折線，根據直觀，我們知道直線就是我們要求的最短折線，以下讓我們用離散變分法的術語表述和求解此一問題。根據「畢達哥拉斯定理」，容易看到上圖中第 n 條直線的長度是 $\sqrt{1 + (\Delta y(n-1))^2}$ 。以上面第 1 條直線為例，由於 $y(0) = 3$ 和 $y(1) = 4$ ，可知該直線的長度為 $\sqrt{1 + (4-3)^2} = \sqrt{2}$ 。根據以上討論，可知上述折線的長度為

$$\sum_{n=x_1+1}^{n=x_2} \sqrt{1 + (\Delta y(n-1))^2} \quad (18)$$

上式就是與「最短折線問題」相關的泛函。把 (18) 與 (16) 比較，可以看到我們有 $F(x, y(x), \Delta y(x-1)) = \sqrt{1 + (\Delta y(n-1))^2}$ ，由此有函數 $F(u, v, w) = \sqrt{1 + w^2}$ 。把此一函數代入 (17) 中等號左端的 F ，可得到

$$\begin{aligned} & \Delta \left(\left[D_w \left(\sqrt{1 + w^2} \right) \right]_{(u,v,w)=(x,y(x),\Delta y(x-1))} \right) - \left[D_v \left(\sqrt{1 + w^2} \right) \right]_{(u,v,w)=(x,y(x),\Delta y(x-1))} \\ &= \Delta \left(\left[\frac{w}{\sqrt{1 + w^2}} \right]_{(u,v,w)=(x,y(x),\Delta y(x-1))} \right) - 0 \\ &= \Delta \left(\frac{\Delta y(x-1)}{\sqrt{1 + (\Delta y(x-1))^2}} \right) \end{aligned}$$

由此可得「最短折線問題」的「歐拉-拉格朗日差分方程」如下：

$$\Delta \left(\frac{\Delta y(x-1)}{\sqrt{1 + (\Delta y(x-1))^2}} \right) = 0, \quad x \in \{(x_1), x_1+1, \dots, x_2-1, (x_2)\}, \quad y(x_1) = y_1, y(x_2) = y_2 \quad (19)$$

從上述方程，容易得到

$$\frac{\Delta y(x-1)}{\sqrt{1 + (\Delta y(x-1))^2}} = c_1 \quad (20)$$

其中 c_1 是任意常數。從上式可以解出 $\Delta y(x-1) = \pm \sqrt{\frac{c_1^2}{1-c_1^2}}$ ，把這個結果中的 $\pm \sqrt{\frac{c_1^2}{1-c_1^2}}$ 換成另一個任意常數 c_2 ，可得

$$\Delta y(x-1) = c_2 \quad (21)$$

上式告訴我們，每條折線的斜率都等於同一個常數 c_2 ，具有此一特徵的折線其實就是直線，由此可知直線就是我們要求的最短折線。請注意我們也可以通過代入 $x-1 = y$ ，然後又代入 $y = x$ 把上式改寫成 $\Delta y(x) = c_2$ 。對此式左右兩端進行不定和分運算，或者利用關係式 $\Delta = E - I$ 把此式改寫成 1 階常係數非齊次差分方程 $Ey(x) - y(x) - c_2 = 0$ ，可求得此一方程的通解為

$$y(x) = c_2x + c_3 \quad (22)$$

上式是直線的方程，把 (19) 中的邊界條件代入上式，便可解出上式中的任意常數 c_2 和 c_3 ，從而得到 (19) 的特解為

$$y(x) = \frac{y_2 - y_1}{x_2 - x_1}x + \frac{x_2y_1 - x_1y_2}{x_2 - x_1}, \quad x \in \{(x_1), x_1 + 1, \dots, x_2 - 1, (x_2)\} \quad (23)$$

接著考慮以下泛函：

$$\sum_{n=1}^{n=10} \left(2 \times \left(\frac{1}{6} \right)^x (y(x))^2 - \left(\frac{1}{6} \right)^{x-1} (\Delta y(x-1))^2 \right) \quad (24)$$

現設要求一個以 $\{(0), 1, \dots, 9, (10)\}$ 為定義域的離散函數 $y(x)$ ，使得上述泛函值為最小，並且滿足 $y(0) = 0, y(10) = 58025$ 。把 (24) 與 (16) 比較，可以看到我們有 $F(x, y(x), \Delta y(x-1)) = 2 \times \left(\frac{1}{6} \right)^x (y(x))^2 - \left(\frac{1}{6} \right)^{x-1} (\Delta y(x-1))^2$ ，由此有函數 $F(u, v, w) = 2 \times \left(\frac{1}{6} \right)^u v^2 - \left(\frac{1}{6} \right)^{u-1} w^2$ 。把此一函數代入 (17) 中等號左端的 F ，可得到

$$\begin{aligned} & \Delta \left(\left[D_w \left(2 \times \left(\frac{1}{6} \right)^u v^2 - \left(\frac{1}{6} \right)^{u-1} w^2 \right) \right]_{(u,v,w)=(x,y(x),\Delta y(x-1))} \right) \\ & - \left[D_v \left(2 \times \left(\frac{1}{6} \right)^u v^2 - \left(\frac{1}{6} \right)^{u-1} w^2 \right) \right]_{(u,v,w)=(x,y(x),\Delta y(x-1))} \\ & = \Delta \left(\left[-2 \times \left(\frac{1}{6} \right)^{u-1} w \right]_{(u,v,w)=(x,y(x),\Delta y(x-1))} \right) - \left[4 \times \left(\frac{1}{6} \right)^u v \right]_{(u,v,w)=(x,y(x),\Delta y(x-1))} \\ & = \Delta \left(-2 \times \left(\frac{1}{6} \right)^{x-1} \Delta y(x-1) \right) - 4 \times \left(\frac{1}{6} \right)^x y(x) \end{aligned}$$

由此可得與 (24) 相關的「歐拉-拉格朗日差分方程」如下：

$$\begin{aligned} & \Delta \left(-2 \times \left(\frac{1}{6} \right)^{x-1} \Delta y(x-1) \right) - 4 \times \left(\frac{1}{6} \right)^x y(x) = 0, \\ & x \in \{(0), 1, \dots, 9, (10)\}, \quad y(0) = 0, y(10) = 58025 \quad (25) \end{aligned}$$

運用 Δ 的定義，然後乘以適當的常數，並對 y 的變項作適當的代入，可以把上式中的方程逐步改寫如下：

$$\begin{aligned}\Delta \left(-2 \times \left(\frac{1}{6} \right)^{x-1} \Delta y(x-1) \right) - 4 \times \left(\frac{1}{6} \right)^x y(x) &= 0 \\ -2 \times \left(\frac{1}{6} \right)^x y(x+1) + \frac{5}{3} \times \left(\frac{1}{6} \right)^{x-1} y(x) - 2 \times \left(\frac{1}{6} \right)^{x-1} y(x-1) &= 0 \\ y(x+1) - 5y(x) + 6y(x-1) &= 0 \\ (E^2 - 5E + 6I)y(x) &= 0\end{aligned}$$

容易求得上述 2 階常係數齊次差分方程的通解為

$$y(x) = c_1 \times 2^x + c_2 \times 3^x \quad (26)$$

把 (25) 中的邊界條件代入上式，便可解出上式中的任意常數 c_1 和 c_2 ，從而得到 (25) 的特解為

$$y(x) = 3^x - 2^x, \quad x \in \{(0), 1, \dots, 9, (10)\} \quad (27)$$

連結至數學專題
連結至周家發網頁